



20 0400 Page 1 of 7 0500

OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/709,020
DATE: 11/27/2000
TIME: 09:23:07

Input Set : A:\ES.txt
Output Set: N:\CRF3\11272000\I709020.raw

ENTERED

3 <110> APPLICANT: Benning, Christoph
4 Sanda, Sherrie
5 Yu, Bin
7 <120> TITLE OF INVENTION: Compositions and Methods for the Synthesis and Subsequent Modification of
8 Uridine-5-Diphosphosulfoquinovose (UDP-SQ)
10 <130> FILE REFERENCE: MSU-04769
C--> 12 <140> CURRENT APPLICATION NUMBER: US/09/709,020
C--> 12 <141> CURRENT FILING DATE: 2000-11-08
12 <160> NUMBER OF SEQ ID NOS: 21
14 <170> SOFTWARE: PatentIn version 3.0
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1134
18 <212> TYPE: DNA
19 <213> ORGANISM: Cyanobacterium Synechococcus
21 <220> FEATURE:
22 <221> NAME/KEY: CDS
23 <222> LOCATION: (1)..(1134)
25 <400> SEQUENCE: 1
26 atg cgc atc gct ctc ttt acc gag acg ttc ctc ccc aaa gtg gat ggc 48
27 Met Arg Ile Ala Leu Phe Thr Glu Thr Phe Leu Pro Lys Val Asp Gly
28 1 5 10 15
30 atc gtc acg cgg ctt cgg cac acg gtc gat cac ctg cag cgt ctt ggc 96
31 Ile Val Thr Arg Leu Arg His Thr Val Asp His Leu Gln Arg Leu Gly
32 20 25 30
34 cac acc gtc atg gtt ttt tgc ccc gac ggc ggg ctc cgc gag cac aag 144
35 His Thr Val Met Val Phe Cys Pro Asp Gly Gly Leu Arg Glu His Lys
36 35 40 45
38 ggg gct cga gtc tat ggg gtt aaa ggc ttt ccg cta ccg ctc tat ccc 192
39 Gly Ala Arg Val Tyr Gly Val Lys Gly Phe Pro Leu Pro Leu Tyr Pro
40 50 55 60
42 gag ctg aag cta gct ttt ccg ttg ccg aaa gtg gga aaa gcc ttg gag 240
43 Glu Leu Lys Leu Ala Phe Pro Leu Pro Lys Val Gly Lys Ala Leu Glu
44 65 70 75 80
46 cgg ttc cgg ccc gac ctg atc cac gtg gtc aat ccg gct gtg ttg ggg 288
47 Arg Phe Arg Pro Asp Leu Ile His Val Val Asn Pro Ala Val Leu Gly
48 85 90 95
50 ttg ggc ggc atc tac tat gcc aag gcg cta aat gtg cca ctc gtg gcg 336
51 Leu Gly Gly Ile Tyr Tyr Ala Lys Ala Leu Asn Val Pro Leu Val Ala
52 100 105 110
54 tcc tat cac acc cat ttg ccg aaa tac ctt gag cat tac ggg ctg ggg 384
55 Ser Tyr His Thr His Leu Pro Lys Tyr Leu Glu His Tyr Gly Leu Gly
56 115 120 125
58 gtc ttg gag ggg gtg ctc tgg gaa ttg ctg aag ctg gcg cat aac caa 432
59 Val Leu Glu Gly Val Leu Trp Glu Leu Leu Lys Leu Ala His Asn Gln
60 130 135 140
62 gca gcg atc aac ctc tgt act tca acc gcg atg gtg cag gag ctg aca 480
63 Ala Ala Ile Asn Leu Cys Thr Ser Thr Ala Met Val Gln Glu Leu Thr

RECEIVED
DEC 14 2000
OIPE/JCWS

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/709,020
 DATE: 11/27/2000
 TIME: 09:23:07

Input Set : A:\ES.txt
 Output Set: N:\CRF3\11272000\I709020.raw

```

64 145          150          155          160
66 gat cac ggc att gag cac tgt tgc ctc tgg cag cga gga glg gat acc      528
67 Asp His Gly Ile Glu His Cys Cys Leu Trp Gln Arg Gly Val Asp Thr
68          165          170          175
70 gag acc ttt cgg cca gac ttg gct act gct gcg atg cgc gat cgc ctc      576
71 Glu Thr Phe Arg Pro Asp Leu Ala Thr Ala Ala Met Arg Asp Arg Leu
72          180          185          190
74 agt ggc ggt aag ccc act gcg ccc ttg ctc tac gtc gga cgc ctc      624
75 Ser Gly Gly Lys Pro Thr Ala Pro Leu Leu Leu Tyr Val Gly Arg Leu
76          195          200          205
78 tca gcc gag aag caa atc gat cgc ctg cga ccc att ttg gat gcc aat      672
79 Ser Ala Glu Lys Gln Ile Asp Arg Leu Arg Pro Ile Leu Asp Ala Asn
80          210          215          220
82 cct gag gct tgc ttg gcc ttg gtc ggc gat gcc ccg cat cgg gcc gaa      720
83 Pro Glu Ala Cys Leu Ala Leu Val Gly Asp Gly Pro His Arg Ala Glu
84 225          230          235          240
86 cta gag caa ttg ttt gct ggc acc cag acg cag ttc att ggc tat ctg      768
87 Leu Glu Gln Leu Phe Ala Gly Thr Gln Thr Gln Phe Ile Gly Tyr Leu
88          245          250          255
90 cat ggg gaa cag cta ggg gcg gcc tac gct tct gct qac gct ttt gtc      816
91 His Gly Glu Gln Leu Gly Ala Ala Tyr Ala Ser Ala Asp Ala Phe Val
92          260          265          270
94 ttt ccc tcc cgg acc gaa acc ctc ggt cta gtc ttg ctg gaa gcc atg      864
95 Phe Pro Ser Arg Thr Glu Thr Leu Gly Leu Val Leu Leu Glu Ala Met
96          275          280          285
98 gca gcg ggt tgt ccg gtc gtg gcg gcc aat tcc ggt gcc att ccc gat      912
99 Ala Ala Gly Cys Pro Val Val Ala Ala Asn Ser Gly Gly Ile Pro Asp
100          290          295          300
102 att gtc agc gac ggc att aat ggt ttc ctg ttc gat cct gag gat gaa      960
103 Ile Val Ser Asp Gly Ile Asn Gly Phe Leu Phe Asp Pro Glu Asp Glu
104 305          310          315          320
106 caa ggg gcg atc gct gcg att cag cgc ttg ttg gct aac cct gca gag      1008
107 Glu Gly Ala Ile Ala Ala Ile Gln Arg Leu Leu Ala Asn Pro Ala Glu
108          325          330          335
110 cgc gag att cta cgc caa gcg gct cgt caa gaa gcc gaa cgc tgg agc      1056
111 Arg Glu Ile Leu Arg Gln Ala Ala Arg Gln Glu Ala Glu Arg Trp Ser
112          340          345          350
114 tgg aac gca gcc acg cgc caa ctc cag gac tac tac tgc gag gtg ttg      1104
115 Trp Asn Ala Ala Thr Arg Gln Leu Gln Asp Tyr Tyr Cys Glu Val Leu
116          355          360          365
118 gca gat ggt tgc tta ccc tta gcg gcc tga      1134
119 Ala Asp Gly Cys Leu Pro Leu Ala Ala
120          370          375
123 <210> SEQ ID NO: 2
124 <211> LENGTH: 377
125 <212> TYPE: PRT
126 <213> ORGANISM: Cyanobacterium Synechococcus
128 <400> SEQUENCE: 2
130 Met Arg Ile Ala Leu Phe Thr Glu Thr Phe Leu Pro Lys Val Asp Gly

```

RAW SEQUENCE LISTING

DATE: 11/27/2000

PATENT APPLICATION: US/09/709,020

TIME: 09:23:07

Input Set : A:\ES.txt

Output Set: N:\CRF3\11272000\I709020.raw

```

131 1          5          10          15
134 Ile Val Thr Arg Leu Arg His Thr Val Asp His Leu Gln Arg Leu Gly
135          20          25          30
138 His Thr Val Met Val Phe Cys Pro Asp Gly Gly Leu Arg Glu His Lys
139          35          40          45
142 Gly Ala Arg Val Tyr Gly Val Lys Gly Phe Pro Leu Pro Leu Tyr Pro
143          50          55          60
146 Glu Leu Lys Leu Ala Phe Pro Leu Pro Lys Val Gly Lys Ala Leu Glu
147 65          70          75          80
150 Arg Phe Arg Pro Asp Leu Ile His Val Val Asn Pro Ala Val Leu Gly
151          85          90          95
154 Leu Gly Gly Ile Tyr Tyr Ala Lys Ala Leu Asn Val Pro Leu Val Ala
155          100          105          110
158 Ser Tyr His Thr His Leu Pro Lys Tyr Leu Glu His Tyr Gly Leu Gly
159          115          120          125
162 Val Leu Glu Gly Val Leu Trp Glu Leu Leu Lys Leu Ala His Asn Gln
163          130          135          140
166 Ala Ala Ile Asn Leu Cys Thr Ser Thr Ala Met Val Gln Glu Leu Thr
167 145          150          155          160
170 Asp His Gly Ile Glu His Cys Cys Leu Trp Gln Arg Gly Val Asp Thr
171          165          170          175
174 Glu Thr Phe Arg Pro Asp Leu Ala Thr Ala Ala Met Arg Asp Arg Leu
175          180          185          190
178 Ser Gly Gly Lys Pro Thr Ala Pro Leu Leu Leu Tyr Val Gly Arg Leu
179          195          200          205
182 Ser Ala Glu Lys Gln Ile Asp Arg Leu Arg Pro Ile Leu Asp Ala Asn
183          210          215          220
186 Pro Glu Ala Cys Leu Ala Leu Val Gly Asp Gly Pro His Arg Ala Glu
187 225          230          235          240
190 Leu Glu Gln Leu Phe Ala Gly Thr Gln Thr Gln Phe Ile Gly Tyr Leu
191          245          250          255
194 His Gly Glu Gln Leu Gly Ala Ala Tyr Ala Ser Ala Asp Ala Phe Val
195          260          265          270
198 Phe Pro Ser Arg Thr Glu Thr Leu Gly Leu Val Leu Leu Glu Ala Met
199          275          280          285
202 Ala Ala Gly Cys Pro Val Val Ala Ala Asn Ser Gly Gly Ile Pro Asp
203          290          295          300
206 Ile Val Ser Asp Gly Ile Asn Gly Phe Leu Phe Asp Pro Glu Asp Glu
207 305          310          315          320
210 Gln Gly Ala Ile Ala Ala Ile Gln Arg Leu Leu Ala Asn Pro Ala Glu
211          325          330          335
214 Arg Glu Ile Leu Arg Gln Ala Ala Arg Gln Glu Ala Glu Arg Trp Ser
215          340          345          350
218 Trp Asn Ala Ala Thr Arg Gln Leu Gln Asp Tyr Tyr Cys Glu Val Leu
219          355          360          365
222 Ala Asp Gly Cys Leu Pro Leu Ala Ala
223          370          375
226 <210> SEQ ID NO: 3
227 <211> LENGTH: 2979

```

RAW SEQUENCE LISTING

DATE: 11/27/2000

PATENT APPLICATION: US/09/709,020

TIME: 09:23:07

Input Set : A:\ES.txt

Output Set: N:\CRF3\11272000\I709020.raw

```

228 <212> TYPE: DNA
229 <213> ORGANISM: Arabidopsis thaliana
231 <400> SEQUENCE: 3
232 ctacacgtta ccttccggta ctggaaacag tegttaate aaccaattga ttggtcccaa 60
234 aacatgaact ttctttttcc tccagaacca aatcgtctga ctgtactgtt cattgcttat 120
236 ctttgtcgtt gctgctctcc aatcatatct ctccgtctct tctcttgcg cttttccaat 180
238 gatctctctt gtttcacggt cgtqcaqtaa agttctcagt ttgttcacgc aatcttcaac 240
240 atctccaggg ttgaacaaaa atccggtttt tccctcctga tgaaaacatc agaatacaga 300
242 aaccacaagc tcaatatagg ttgaccata agaacaatca atgcaagatc attttgtgta 360
244 ccagctctatg attgaataaa gtttcagtcc ggttacagct cgtttataag aaaattggca 420
246 gaaattgttt ttcaaccat ttccgttcgg ttgatatgct catcaatatg gtttggcagt 480
248 taattqtaat tcagataaatt cactgacctg atcttcaggg attatatcag qgattccacc 540
250 ggcacgggccc gcgacgacag gaagtctctg agacattgct tcaagaacca caaagccaa 600
252 tgtctccgac tctgatggca tcacaaacac atctccactt gcgtaagctt gtgagagtcc 660
254 ategccttgt aacgttccag tgaaaaccgc tggcattccg gtaaaacact tctcaagatc 720
256 ctctctttta gaaaaacaaa cagataaaca aaattacaat gttgttgaat agaaatcttc 780
258 agataacaaat atggccaatc tttacaaaaa ctagtacttg tatgtccat ctccaatgaa 840
260 agcaalcgca gcttcaggta atttgtccat taccctgcac acaaaattct caatatcaaa 900
262 attcgataca ccacttaaaa gaagtgaatc cagtttatac aaaattctaa cctctttaa 960
264 agctccaaac tctttctctc gccaatgcga cctacatgaa tcaactagtg cttttctggt 1020
266 tgcctattac tgttaattac aaatatatta acatcaagat tagcgtggaa agtatcattg 1080
268 tttttaatgc atataaaaga aacgtatatt ctattcttgc ctcagtctta taagcatttc 1140
270 ttgagaacgg aaacggggat tgaagcttcc tgaatcgaca cctttattcc aaagtcaag 1200
272 ttgattagct tcaatatatg aaaaaaagaa gaagaaaatg taagttttga acaatcatag 1260
274 agcttgtaaa aaaatgtaat gtatgatctt tcttcacctg cagttgcacc agctgctata 1320
276 agatcttttc caatggcagc agaaggaaat aatgtaagat cagccgtctt gtgaaggaa 1380
278 cctgaataaa gcataatcaa gtttagtttc atattatata tcacaataaa ccagaaaaag 1440
280 aagaaggaaa ttttgacatt tgaaaagcgg gttttacata cttattatag accacattgg 1500
282 ttttaccaac caactaaaag tgtatcttgg gatgtatctg cacatcaaca caagcttaat 1560
284 cttagacaaa aattttttta taacaacatt gtgaaatgag gcagaaaaag gtacttacac 1620
286 agggacgtgt gtgtggttaag acattactat tggtaacagat agcatttttg ctattgccag 1680
288 agcaccaaaq accataaaat tttagtccag gaaaaaagag agtcaagttc tggattctct 1740
290 ccagttcact ggtcttctga atgttttagt aatgtgaatt cttgagggat ttaccataac 1800
292 tccgggagat gaagcgtgta taatgtcagg cttaaaccgt gcaatttcag agatgattct 1860
294 tggactaagc gcaagcgaga gtggaacctt ttggttaataa ggacaaggga agctacagaa 1920
296 gagaagaaga aattagcgat attaccaaat agagaacatc cagtgaagta actaaatggt 1980
298 gctaccttct tgatccaatg actctggctc cataaaactc ttcaggaaaca ccttcatgtg 2040
300 tgcgcacgac tataacctat gaagcaaaaa agttatttaa aaaaaaaa aaggaacagt 2100
302 taacacttgt caagtaattc taattctgga aacagttact tatgagctga ctgaaaagat 2160
304 acttaagtgt aagaatgaga tagtataaga agaaacctcg tctccattt cagggaggta 2220
306 tctaataaaa tcttggaaat tgtttttgta gccggataca tagctgcaac aaaatcaag 2280
308 agagaatcac ttccataaat aacatgacat ataataaag cttttggta atggatcgtt 2340
310 gattccgaga atcttgggat attcacaact aaaatctgac aactttgact caaacaatc 2400
312 ctgaatgtaa ttggttttaa cgtctacta tataatttgc taaattggtg gtgtagcaaa 2460
314 ttcatacatt agcaggtatc tcttcataaa ataaatgtaa cgtacaaatc gaagaaaaa 2520
316 aacattacag gaaaaagtca accaaggaaa aaatgagta gaattctgtt tcacagagac 2580
318 atttctcga acacaaaaa agcaaaaaa gaacactgtg aagaagactt acgcaaggg 2640
320 aagaaggctc acaagagag caattctctt aggtttagag agcagactcag gatcgagaag 2700
322 cggcgcatca atctccgatt cytcatttcc tctgacttga gtaatagtca tatcgttgg 2760

```

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/709,020 DATE: 11/27/2000
 TIME: 09:23:08

Input Set : A:\ES.txt
 Output Set: N:\CRF3\11272000\I709020.raw

```

324 cccagaaaaca gcttcccttgg taatcactcc acagaagcgg agcttgcttt tcttgctaatt 2820
326 gggtaaatctc cggtagccaa acgaaagagg ggaatgaaga acaaaagaag aagacctggg 2880
328 aggagaacaa gaggttgccg aggaagaaga acaagtgtta gtcgtgctag gaagcaaatg 2940
330 aggagggtata gagagatttta tagaagaaag agtcgtcat 2979
333 <210> SEQ ID NO: 4
334 <211> LENGTH: 2537
335 <212> TYPE: DNA
336 <213> ORGANISM: Arabidopsis thaliana
338 <400> SEQUENCE: 4
339 tcataatcttg aaaaagcaact tcttcttctt cttcttcttc ttcttcttct tcttcttccc 60
341 atcaaaaagt gctagtggca agaccggcca tattccagta acaaaagttta aattcaaaaac 120
343 tatggaacac aatttgctcc atgttttggt gccatgaatta aacgaccata tcacttggtt 180
345 cggccacttc atctcggata cgcacaagcg gttgtcaagg ttgtaaaaga cgattttgta 240
347 aggatctaca ttggcaagg gagctttaca gacgacttga aaagtctcgg tgtgaagatt 300
349 gaaaaggtaaa actttgggttt ctttgcactc ggtgaaccaa taagcgacc catctagata 360
351 cacaggtgag ggttaagcag caatccgata aggagcagcg agagtgcac accctcaagc 420
353 attggtacta aagtcgaaaa cttcgcctgt agtagcgttt tttaggccta ttctgcaga 480
355 gtltgtatac caaacgggct tgtatgtgce cctgaatttg tctttaccga atccaagcat 540
357 aaagaatttg cgttttaagct tgtagtaacc atctcgtaag tcgatacatga gtttttgata 600
359 atcgcacaaa ggaagagggc gataccatct agtggtcggg ttgaccacat aaccggattt 660
361 gtcgtgattg taagacaaa cgaagccgtc acaactacta tgtgaaacta agtacagtac 720
363 gttatccttc tcccaaggag tagggatctt gaacgttgat gatgaaccca actgcaatgt 780
365 tcttaagat tctatggctg ggtttataac atggtgagga aatacagaca ccctcagaac 840
367 atctggatct cctgattgct gacgatgttt caactaaaa tgtatacat tttattttat 900
369 taaaagata aatatcattt tgttgaatt ggaacacaaa atctaaaaga taacttttaga 960
371 tatgttaagg tcgcataatg atttgattac attgctatga gtatatgact cacttttcca 1020
373 ataataaaaa aaataaagat gttgagctct actaat taag ggtaaatata aaatttctga 1080
375 tcattaatat aaagaaaaag cctctcataa gccaaagcca ttttactcgc cgtaaacatc 1140
377 tccgcagcat actctttaca cttccttctt cctcgcgcca atctctcgc tcttccagcc 1200
379 acagcaacct ccataaccgc cgtcaaaagt tccacatttg gcgcaaacat aaaccacac 1260
381 tcatattcca ccaactatgt cctctttatg cttgcgtacc tagatgccc cactggttta 1320
383 ccactcaaca tcgcttccat taacgtttaa tcaagacctt gtygtctaaq cgttgatttc 1380
385 acgaacaaat cgtatccgtt gtagaaaccc ttgagctcat ttgggtttag agatcccaaa 1440
387 atggaaactt tttcccttaa ttcttgttaa cgttgctccc atggccctga tccagctact 1500
389 acgaggtaaa cat.ttgata cgtttggatt attttcgcga aagcttcgaa gagcaatgga 1560
391 tgtcctttgt ctttgactaa tctcccagca gctcctaaaa caatcgctga tgagttttct 1620
393 ggtaacccca attttgacct aaacagagta cgtagcttct tgtctgatgt gaatccgttc 1680
395 tcgtcgactc catttgaggat cacatgaacc cttttctcag ggaatttgta aacgtctctt 1740
397 agcat.ttct cgcagctatc gctgatagcg atgtggtgag cgtagt.tgt gaagaatctg 1800
399 atttcgtcca gtatcttggg aagcacggca ccgtatagac ttgcattgaa gccttgtgat 1860
401 ct.tggttcgt ctggtttacg gatcaggtct tggtaaatat ttgattgtaa gctctctaac 1920
403 gcaatgcctt gccaggtatc agcgaggttt ggaacctccc gggcgatcca gtgaggtaaa 1980
405 gcaacacttt cagagtgaac cgcacgaaa ggttctttct tgttttcttc ttggtaaaagc 2040
407 tcccatgctt tgtt.tgaccg ccattttccg ggtcccgct cgcctgaga atggattata 2100
409 gggtaaatga tt.tggtcggg aacagggggg attttgttgg tt.tcagggga ttggtctaag 2160
411 ggagaggtga aaacgtggac acggtgtcca cggcgagcta aggcggtgta tagagtgaag 2220
413 gcgtggcgtt ccattgccgc aggattggga cccgtgggac atttcgggga gaaaacagcg 2280
415 agtttttagt ttttgggagg aggg.tggtt aaggagaaat caaggcgtt ccaggcaaat 2340
417 tgggcggttt ggaggtcgcc agaccacggt ggcgtggttt tgtcagagga ggaggaggag 2400

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/709,020

DATE: 11/27/2000

TIME: 09:23:09

Input Set : A:\ES.txt

Output Set: N:\CRF3\11272000\I709020.raw

L:12 M:270 C: Current Application Number differs, Replaced Current Application No
L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date